

SEQUENCE LISTING



#4

(1) GENERAL INFORMATION:

(i) APPLICANT: Somerville, Chris
Broun, Pierre
van de Loo, Frank

(ii) TITLE OF INVENTION: Production of Hydroxylated Fatty Acids in Genetically Modified Plants

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: CUSHMAN DARBY & CUSHMAN, LLP
(B) STREET: 1100 NEW YORK AVENUE, N.W.
(C) CITY: WASHINGTON
(D) STATE: D.C.
(E) COUNTRY: USA
(F) ZIP: 20005-3918

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
(B) COMPUTER: IBM PC
(C) OPERATING SYSTEM: DOS 5.0
(D) SOFTWARE: Word Perfect 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/530,862
(B) FILING DATE: September 20, 1995
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Paul N. Kokulis
(B) REGISTRATION NUMBER: 16,773
(C) REFERENCE/DOCKET NUMBER: 1220/213781

(viii) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 202 861 3000
(B) TELEFAX: 202 822 0944

(2) INFORMATION FOR SEQ ID NO:1

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 543 nucleotides
(B) TYPE: nucleotide
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATTGGCACC GCGGGCACCA TTCCAACAAT GGATCCCTAG AAAAAGATGA AGTCTTTGTC	60
CCACCTAAGA AAGCTGCAGT CANATGGTAT GTCAAATACC TCAACAACCC TCTTGGACGC	120
ATTCTGGTGT TAACAGTTCA GTTTATCCTC GGGTGGCCTT TGTATCTAGC CTTTAATGTA	180
TCAGGTAGAC CTTATGATGG TTTCGCTTCA CATTTCTTCC CTCATGCACC TATCTTTAAG	240
GACCGTGAAC GTCTCCAGAT ATACATCTCA GATGCTGGTA TTCTAGCTGT CTGTTATGGT	300
CTTTACCGTT ACGCTGCTTC ACAAGGATTG ACTGCTATGA TCTGCGTCTA CGGAGTACCG	360

CTTTTGATAG TGAAC TTTT CCTTGTCTTG GTCAC TTTCT TGCAGCACAC TCATCCTTCA	420
TTACCTCACT ATGATTCAAC CGAGTGGGAA TGGATTAGAG GAGCTTTGGT TACGGTAGAC	480
AGAGACTATG GAATCTTGAA CAAGGTGTTT CACAACATAA CAGACACCCA CGTAGCACAC	540
CAC	543

(2) INFORMATION FOR SEQ ID NO:2

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	544 nucleotides
(B) TYPE:	nucleotide
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TATAGGCACC GGAGGCACCA TTCCAACACA GGATCCCTCG AAAGAGATGA AGTATTTGTC	60
CCAAAGCAGA AATCCGCAAT CAAGTGGTAC GGCGAATACC TCAACAACCC TCCTGGTCGC	120
ATCATGATGT TAACTGTCCA GTTCGTCCTC GGATGGCCCT TGTACTTAGC CTTCAACGTT	180
TCTGGCAGAC CCTACAATGG TTTCGCTTCC CATTTCTTCC CCAATGCTCC TATCTACAAC	240
GACCGTGAAC GCCTCCAGAT TTACATCTCT GATGCTGGTA TTCTAGCCGT CTGTTATGGT	300
CTTTACCGTT ACGCTGTTGC ACAAGGACTA GCCTCAATGA TCTGTCTAAA CGGAGTTCCG	360
CTTCTGATAG TTAAC TTTT CCTCGTCTTG ATCACTTACT TACAACACAC TCACCCTGCG	420
TTGCCTCACT ATGATTCATC AGAGTGGGAT TGGCTTAGAG GAGCTTTAGC TACTGTAGAC	480
AGAGACTATG GAATCTTGAA CAAGGTGTTT CATAACATCA CAGACACCCA CGTCGCACAC	540
CACT	544

(2) INFORMATION FOR SEQ ID NO:3

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	1855 nucleotides
(B) TYPE:	nucleotide
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

(ii) MOLECULE TYPE: genomic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAAGCTTT ATAAGAAGTT AGTTTTCTCT GGTGACAGAG AAATTNTGTC AATTGGTAGT	60
GACAGTTGAA GCAACAGGAA CAACAAGGAT GGTGTTGNT GATGCTGATG TGGTGATGTG	120
TTATTCATCA AATACTAAAT ACTACATTAC TTGTTGCTGC CTACTTCTCC TATTTCTCTCC	180
GCCACCCATT TTGGACCCAC GANCCTTCCA TTAAACCCT CTCTCGTGCT ATTCACCAGA	240
AGAGAAGCCA AGAGAGAGAG AGAGAGAATG TTCTGAGGAT CATTGTCTTC TTCATCGTTA	300

TTAACGTAAG TTTTTTTTGA CCACTCATAT CTAAAATCTA GTACATGCAA TAGATTAATG	360
ACTGTTCCCTT CTTTTGATAT TTTCAGCTTC TTGAATTCAA GATGGGTGCT GGTGGAAGAA	420
TAATGGTTAC CCCCTCTTCC AAGAAATCAG AAAGTGAAGC CCTAAAACGT GGACCATGTG	480
AGAAACCACC ATTCAGTGT AAAGATCTGA AGAAAGCAAT CCCACAGCAT TGTTTCAAGC	540
GCTCTATCCC TCGTTCTTTC TCCTACCTTC TCACAGATAT CACTTTAGTT TCTTGCTTCT	600
ACTACGTGTC CACAAATTAC TTCTCTCTTC TTCCTCAGCC TCTCTCTACT TACCTAGCTT	660
GGCCTCTCTA TTGGGTATGT CAAGGCTGTG TCTTAACCGG TATCTGGGTC ATTGGCCATG	720
AATGTGGTCA CCATGCATTC AGTGACTATC AATGGGTAGA TGACACTGTT GGTTTTATCT	780
TCCATTCCTT CCTTCTCGTC CCTTACTTCT CCTGGAAATA CAGTCATCGT CGTCACCATT	840
CCAACAATGG ATCTCTCGAG AAAGATGAAG TCTTTGTCCC ACCGAAGAAA GCTGCAGTCA	900
AATGGTATGT TAAATACCTC AACAACCCTC TTGGACGCAT TCTGGTGTTA ACAGTTCAGT	960
TTATCCTCGG GTGGCCTTTG TATCTAGCCT TTAATGTATC AGGTAGACCT TATGATGGTT	1020
TCGCTTCACA TTTCTTCCCT CATGCACCTA TCTTTAAAGA CCGAGAACGC CTCCAGATAT	1080
ACATCTCAGA TGCTGGTATT CTAGCTGTCT GTTATGGTCT TTACCGTTAC GCTGCTTCAC	1140
AAGGATTGAC TGCTATGATC TCGTCTATG GAGTACCGCT TTTGATAGTG AACTTTTTCC	1200
TTGTCTTGGT AACTTTCTTG CAGCACACTC ATCCTTCGTT ACCTCATTAT GATTCAACCG	1260
AGTGGGAATG GATTAGAGGA GCTTTGGTTA CGGTAGACAG AGACTATGGA ATATTGAACA	1320
AGGTGTTCCA TAACATAACA GACACACATG TGGCTCATCA TCTCTTTGCA ACTATACCGC	1380
ATTATAACGC AATGGAAGCT ACAGAGGCGA TAAAGCCAAT ACTTGGTGAT TACTACCACT	1440
TCGATGGAAC ACCGTGGTAT GTGGCCATGT ATAGGGAAGC AAAGGAGTGT CTCTATGTAG	1500
AACCGGATAC GGAACGTGGG AAGAAAGGTG TCTACTATTA CAACAATAAG TTATGAGGCT	1560
GATAGGGCGA GAGAAGTGCA ATTATCAATC TTCATTTCCA TGTTTTAGGT GTCTTGTTTA	1620
AGAAGCTATG CTTTGTTTCA ATAATCTCAG AGTCCATNTA GTTGTGTTCT GGTGCATTTT	1680
GCCTAGTTAT GTGGTGTCGG AAGTTAGTGT TCAAAGTCT TCCTGCTGTG CTGCCCAGTG	1740
AAGAACAAGT TTACGTGTTT AAAATACTCG GAACGAATTG ACCACAANAT ATCCAAAACC	1800
GGCTATCCGA ATTCCATATC CGAAAACCGG ATATCCAAAT TTCCAGAGTA CTTAG	1855

(2) INFORMATION FOR SEQ ID NO:4

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 384 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Ala Gly Gly Arg Ile Met Val Thr Pro Ser Ser Lys Lys Ser
 10
 Glu Thr Glu Ala Leu Lys Arg Gly Pro Cys Glu Lys Pro Pro Phe Thr
 20 30
 Val Lys Asp Leu Lys Lys Ala Ile Pro Gln His Cys Phe Lys Arg Ser
 40
 Ile Pro Arg Ser Phe Ser Tyr Leu Leu Thr Asp Ile Thr Leu Val Ser
 50 60
 Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
 70 80
 Leu Ser Thr Tyr Leu Ala Trp Pro Leu Tyr Trp Val Cys Gln Gly Cys
 90
 Val Leu Thr Gly Ile Trp Val Ile Gly His Glu Cys Gly His His Ala
 100 110
 Phe Ser Asp Tyr Gln Trp Val Asp Asp Thr Val Gly Phe Ile Phe His
 120
 Ser Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg
 130 140
 His His Ser Asn Asn Gly Ser Leu Glu Lys Asp Glu Val Phe Val Pro
 150 160
 Pro Lys Lys Ala Ala Val Lys Trp Tyr Val Lys Tyr Leu Asn Asn Pro
 170
 Leu Gly Arg Ile Leu Val Leu Thr Val Gln Phe Ile Leu Gly Trp Pro
 180 190
 Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala
 200
 Ser His Phe Phe Pro His Ala Pro Ile Phe Lys Asp Arg Glu Arg Leu
 210 220
 Gln Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu
 230 240
 Tyr Arg Tyr Ala Ala Ser Gln Gly Leu Thr Ala Met Ile Cys Val Tyr
 250
 Gly Val Pro Leu Leu Ile Val Asn Phe Phe Leu Val Leu Val Thr Phe
 260 270
 Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Thr Glu Trp
 280
 Glu Trp Ile Arg Gly Ala Leu Val Thr Val Asp Arg Asp Tyr Gly Ile
 290 300
 Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His
 310 320
 Leu Phe Ala Thr Ile Pro His Tyr Asn Ala Met Glu Ala Thr Glu Ala
 330
 Ile Lys Pro Ile Leu Gly Asp Tyr Tyr His Phe Asp Gly Thr Pro Trp
 340 350

Tyr Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Leu Tyr Val Glu Pro
360

Asp Thr Glu Arg Gly Lys Lys Gly Val Tyr Tyr Tyr Asn Asn Lys Leu
370 380

(2) INFORMATION FOR SEQ ID NO:5

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 387 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gly Gly Gly Gly Arg Met Ser Thr Val Ile Thr Ser Asn Asn Ser
10

Glu Lys Lys Gly Gly Ser Ser His Leu Lys Arg Ala Pro His Thr Lys
20 30

Pro Pro Phe Thr Leu Gly Asp Leu Lys Arg Ala Ile Pro Pro His Cys
40

Phe Glu Arg Ser Phe Val Arg Ser Phe Ser Tyr Val Ala Tyr Asp Val
50 60

Cys Leu Ser Phe Leu Phe Tyr Ser Ile Ala Thr Asn Phe Phe Pro Tyr
70 80

Ile Ser Ser Pro Leu Ser Tyr Val Ala Trp Leu Val Tyr Trp Leu Phe
90

Gln Gly Cys Ile Leu Thr Gly Leu Trp Val Ile Gly His Glu Cys Gly
100 110

His His Ala Phe Ser Glu Tyr Gln Leu Ala Asp Asp Ile Val Gly Leu
120

Ile Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser
130 140

His Arg Arg His His Ser Asn Ile Gly Ser Leu Glu Arg Asp Glu Val
150 160

Phe Val Pro Lys Ser Lys Ser Lys Ile Ser Trp Tyr Ser Lys Tyr Ser
170

Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Ala Ala Thr Leu Leu Leu
180 190

Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
200

Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Phe Ser Glu Arg
210 220

Glu Arg Leu Gln Ile Tyr Ile Ala Asp Leu Gly Ile Phe Ala Thr Thr
230 240

Phe Val Leu Tyr Gln Ala Thr Met Ala Lys Gly Leu Ala Trp Val Met
 250
 Arg Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Cys Phe Leu Val Met
 260 270
 Ile Thr Tyr Leu Gln His Thr His Pro Ala Ile Pro Arg Tyr Gly Ser
 280
 Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg Asp
 290 300
 Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His Val
 310 320
 Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu Ala
 330
 Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp Gly
 340 350
 Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu Phe
 360
 Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr Arg
 370 380
 Asn Lys Tyr

(2) INFORMATION FOR SEQ ID NO:6

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 383 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
 10
 Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
 20 30
 Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
 40
 Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
 50 60
 Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
 70 80
 Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
 90
 Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
 100 110

Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
 120
 Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
 130 140
 His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
 150 160
 Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
 170
 Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu
 180 190
 Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys
 200
 His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
 210 220
 Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
 230 240
 Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly
 250
 Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu
 260 270
 Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
 280
 Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
 290 300
 Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
 310 320
 Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile
 330
 Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr
 340 350
 Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
 360
 Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu
 370 380

(2) INFORMATION FOR SEQ ID NO:7

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 384 amino acids.
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Gly Ala Gly Gly Arg Met Gln Val Ser Pro Pro Ser Lys Lys Ser
 10
 Glu Thr Asp Asn Ile Lys Arg Val Pro Cys Glu Thr Pro Pro Phe Thr
 20 30
 Val Gly Glu Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
 40
 Ile Pro Arg Ser Phe Ser His Leu Ile Trp Asp Ile Ile Ile Ala Ser
 50 60
 Cys Phe Tyr Tyr Val Ala Thr Thr Tyr Phe Pro Leu Leu Pro Asn Pro
 70 80
 Leu Ser Tyr Phe Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
 90
 Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His Ala Ala Phe
 100 110
 Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
 120
 Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
 130 140
 His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Arg
 150 160
 Arg Ser Gln Thr Ser Ser Gly Thr Ala Ser Thr Ser Thr Thr Phe Gly
 170
 Arg Thr Val Met Leu Thr Val Gln Phe Thr Leu Gly Trp Pro Leu Tyr
 180 190
 Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Gly Phe Ala Cys
 200
 His Phe His Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
 210 220
 Ile Tyr Ile Ser Asp Ala Gly Ile Leu Ala Val Cys Tyr Gly Leu Leu
 230 240
 Pro Tyr Ala Ala Val Gln Gly Val Ala Ser Met Val Cys Phe Leu Arg
 250
 Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile Thr Tyr Leu
 260 270
 Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
 280
 Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
 290 300
 Asn Gln Gly Phe His Asn Ile Thr Asp Thr His Glu Ala His His Leu
 310 320
 Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile
 330
 Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly Thr Pro Val Val
 340 350

Lys Ala Met Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
360

Arg Gln Gly Glu Lys Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu Xaa
370 380

(2) INFORMATION FOR SEQ ID NO:8

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 309 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ser Leu Leu Thr Ser Phe Ser Tyr Val Val Tyr Asp Leu Ser Phe Ala
10

Phe Ile Phe Tyr Ile Ala Thr Thr Tyr Phe His Leu Leu Pro Gln Pro
20 30

Phe Ser Leu Ile Ala Trp Pro Ile Tyr Trp Val Leu Gln Gly Cys Leu
40

Leu Thr Arg Val Cys Gly His His Ala Phe Ser Lys Tyr Gln Trp Val
50 60

Asp Asp Val Val Gly Leu Thr Leu His Ser Thr Leu Leu Val Pro Tyr
70 80

Phe Ser Trp Lys Ile Ser His Arg Arg His His Ser Asn Thr Gly Ser
90

Leu Asp Arg Asp Glu Arg Val Lys Val Ala Trp Phe Ser Lys Tyr Leu
100 110

Asn Asn Pro Leu Gly Arg Ala Val Ser Leu Leu Val Thr Leu Thr Ile
120

Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
130 140

Ser Phe Ala Ser His Tyr His Pro Tyr Arg Val Arg Leu Leu Ile Tyr
150 160

Val Ser Asp Val Ala Leu Phe Ser Val Thr Tyr Ser Leu Tyr Arg Val
170

Ala Thr Leu Lys Gly Leu Val Trp Leu Leu Cys Val Tyr Gly Val Pro
180 190

Leu Leu Ile Val Asn Gly Phe Leu Val Thr Ile Thr Tyr Leu Arg Val
200

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala Thr
210 220

Met Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His His Ile Thr
230 240

Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His
250

Leu Arg Val Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Asp Thr
260 270

Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Arg Glu Cys Leu Tyr Val
280

Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr Trp Tyr Arg Asn
290 300

Lys Tyr Leu Arg Val

(2) INFORMATION FOR SEQ ID NO:9

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 302 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile Ala Phe Cys Leu Tyr Tyr
10

Val Ala Thr His Tyr Phe His Leu Leu Pro Gly Pro Leu Ser Phe Arg
20 30

Gly Met Ala Ile Tyr Trp Ala Val Gln Gly Cys Ile Leu Thr Gly Val
40

Trp Val Val Ala Phe Ser Asp Tyr Gln Leu Leu Asp Asp Ile Val Gly
50 60

Leu Ile Leu His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr
70 80

Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu
90

Val Phe Val Pro Lys Val Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg
100 110

Val Leu Thr Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu
120

Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr
130 140

Asp Pro Tyr Gly Pro Ile Tyr Ser Val Ile Ser Asp Ala Gly Val Leu
150 160

Ala Val Val Tyr Gly Leu Phe Arg Leu Ala Met Ala Lys Gly Leu Ala
170

Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu Val Val Asn Gly Phe
180 190

Leu Val Leu Ile Thr Phe Leu Gln His Thr His Val Ser Glu Trp Asp
 200
 Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
 210 220
 Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
 230 240
 Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Val Glu Tyr
 250
 Tyr Arg Phe Asp Glu Thr Pro Phe Val Lys Ala Met Trp Arg Glu Ala
 260 270
 Arg Glu Cys Ile Tyr Val Glu Pro Asp Gln Ser Thr Glu Ser Lys Gly
 280
 Val Phe Trp Tyr Asn Asn Lys Leu Ala Met Glu Ala Thr Val
 290 300

(2) INFORMATION FOR SEQ ID NO:10

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 372 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
 10
 Gln Leu Ala Arg Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val
 20 30
 Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
 40
 His Cys Phe Glu Arg Ser Val Leu Lys Ser Phe Ser Tyr Val Val His
 50 60
 Asp Leu Val Ile Ala Ala Ala Leu Leu Tyr Phe Ala Leu Ala Ile Ile
 70 80
 Pro Ala Leu Pro Ser Pro Leu Arg Tyr Ala Ala Trp Pro Leu Tyr Trp
 90
 Ile Ala Gln Gly Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Val Val
 100 110
 Gly Leu Val Leu His Ser Ser Leu Met Val Pro Tyr Phe Ser Trp Lys
 120
 Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp
 130 140
 Glu Val Phe Val Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro
 150 160

Tyr Val Tyr Asn Asn Pro Val Gly Arg Val Val His Ile Val Val Gln
 170
 Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg
 180 190
 Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
 200
 Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Val
 210 220
 Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp
 230 240
 Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp
 250
 Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His
 260 270
 Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
 280
 Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
 290 300
 Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala
 310 320
 Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His
 330
 Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu
 340 350
 Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn
 360
 Lys Lys Phe Xaa
 370

(2) INFORMATION FOR SEQ ID NO:11

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 224 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Trp Val Met Ala His Asp Cys Gly His His Ala Phe Ser Asp Tyr Gln
 10
 Leu Leu Asp Asp Val Val Gly Leu Ile Leu His Ser Cys Leu Leu Val
 20 30
 Pro Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr
 40

Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Lys Lys Ser Ser
 50 60
 Ile Arg Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Ile Met
 70 80
 Thr Ile Ala Val Thr Leu Ser Leu Gly Trp Pro Leu Tyr Leu Ala Phe
 90
 Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro
 100 110
 Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ile Glu Ile Phe Ile Ser
 120
 Asp Ala Gly Val Leu Ala Val Thr Phe Gly Leu Tyr Gln Leu Ala Ile
 130 140
 Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu
 150 160
 Val Val Asn Ser Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His
 170
 Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly
 180 190
 Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe
 200
 His Asn Ile Thr Asp Thr Gln Val Ala His His Leu Phe Thr Met Pro
 210 220

(2) INFORMATION FOR SEQ ID NO:12

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 nucleotides
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCTCTTTTGT GCGCTCATTC

20

(2) INFORMATION FOR SEQ ID NO:13

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 nucleotides
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGTACCAGA AAACGCCTTG

20

(2) INFORMATION FOR SEQ ID NO:14

(i) SEQUENCE CHARACTERISTICS:

(A)	LENGTH:	20 nucleotides
(B)	TYPE:	nucleotide
(C)	STRANDEDNESS:	single
(D)	TOPOLOGY:	linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TAYWSNCAYM GNMGNCA YCA

20

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

(A)	LENGTH:	21 nucleotides
(B)	TYPE:	nucleotide
(C)	STRANDEDNESS:	single
(D)	TOPOLOGY:	linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

RTGRTGNGCN ACRTGNGTRT C

21